

mgr Victor Murcia Pienkowski

**Mapowanie punktów złamań
w zrównoważonych objawowych translokacjach chromosomowych**

**Rozprawa na stopień doktora nauk medycznych i nauk o zdrowiu
w dyscyplinie nauki medyczne**

Promotor: prof. dr hab. n. med. Rafał Płoski

**Zakład Genetyki Medycznej, Studium Medycyny Molekularnej,
Warszawski Uniwersytet Medyczny**



**Obrona rozprawy doktorskiej przed Radą Dyscypliny Nauk Medycznych
Warszawskiego Uniwersytetu Medycznego**

Warszawa 2020

1. STRESZCZENIE

Zrównoważone translokacje chromosomowe (ang. balanced chromosomal translocations, BCTs) to stosunkowo często obserwowane strukturalne aberracje chromosomowe. Większość nosicieli BCTs pozostaje bezobjawowa, ale w 26,8% przypadków BCTs związany jest z patologicznym obrazem klinicznym u pacjenta. *De novo* BCTs zdarzają się raz na 2000 żywych urodzeń. W związku z tak wysokim odsetkiem symptomatycznych pacjentów wśród nosicieli BCTs, nasuwa się przyczynowe powiązanie rearanżacji z obserwowanym fenotypem. Nosiciele objawowych BCTs są zatem naturalnym układem eksperymentalnym, w którym uszkodzone geny i/lub elementy regulatorowe odpowiadające za fenotyp zlokalizowane są w punktach złamania. Przynajmniej trzy mechanizmy molekularne mogą tłumaczyć związek BCTs z fenotypem klinicznym: (i) bezpośrednie uszkodzenie genu skutkujące zaburzeniem ekspresji, (ii) powstawanie nowych genów lub nowych połączeń promotor-gen w wyniku fuzji genów, oraz (iii) delokalizacja i/lub uszkodzenie elementów regulatorowych wpływające na ekspresję nienaruszonych genów znajdujących się w pobliżu miejsca złamania – efekt pozycji. Punkty złamań BCTs są z reguły mapowane z użyciem konwencjonalnych metod takich jak, kariotypowanie, fluorescencyjna hybrydyzacja in situ (FISH), hybrydyzacja typu Southern blot czy PCR długich fragmentów. Wszystkie są pracochłonne i czasochłonne, a uzyskane wyniki mają niską rozdzielczość w punkcie złamania.

Celem powyższej pracy była identyfikacja nowych genów/*loci*, będących przyczyną choroby u objawowych nosicieli *de novo* BCTs.

Z badania wykluczono pacjentów, u których:

- wykryta zrównoważona translokacja jest translokacją Robertsonowską.
- wykryta zrównoważona translokacja jest odziedziczona od któregoś z rodziców.
- wykryto w aCGH zmianę liczby kopii (delecja lub amplifikacja) w punktach złamania chromosomu lub w innych regionach genomu, która nie była obecna u żadnego z rodziców.
- punkt złamania w wykrytej translokacji chromosomowej znajduje się w obrębie centromeru.

Do badania zakwalifikowano 11 pacjentów z BCTs. Do identyfikacji dokładnej struktury translokacji wykorzystano całogenomowe sekwencjonowanie typu „mate pair” w technologii sekwencjonowania następnej generacji. Wyniki potwierdzano z wykorzystaniem „złotego standardu”, którym jest sekwencjonowanie Sangera. W przypadku, gdy translokacja nie

uszkadzała bezpośrednio genu, weryfikowano ekspresję genów, zlokalizowanych w funkcjonalnej domenie uszkodzonej przez translokację, za pomocą qPCR.

U wszystkich 11 analizowanych pacjentów zidentyfikowano dokładną strukturę punktów złamań BCTs. U pięciu probantów translokacja uszkadzała geny o nieznanym wpływie na ludzki fenotyp. U jednego pacjenta translokacja na obu chromosomach znajdowała się w niekodującym DNA. U pozostałych pacjentów uszkodzone zostały geny o znanym wpływie na ludzki fenotyp. Do genów, których uszkodzenie nie zostało opisane u ludzi zaliczają się: *BAHD1*, *EFNA5*, *PPP2R5E*, *SLC4A10* oraz *TXNDC5*. Prawdopodobnie wszystkie te geny, z wyjątkiem *TXNDC5*, są odpowiedzialne za fenotyp obserwowany u pacjentów. W przypadku *TXNDC5* choroba raczej jest związana z prawdopodobną zmianą ekspresji genu *FOXG1* znajdującego się w uszkodzonej domenie. Translokacja znajdująca się w niekodującym DNA skutkowałą podwyższeniem ekspresji genu *DKK2* prawdopodobnie odpowiedzialnego za chorobę pacjentki. Uszkodzone geny *EBF3*, *NAA15*, *NIPBL*, *RET*, *SLC6A1* oraz *TRPS1* mają znany wpływ na fenotyp u ludzi oraz są przyczyną patologii występujących u pacjentów. Status pacjenta, u którego doszło do złamania genu *ZNF423* nie jest oczywisty w związku z nietypowym fenotypem probanta w porównaniu do innych osób, u których wykryto mutacje dominujące w tym genie.

Podsumowując, wykorzystanie sekwencjonowania całogenomowego typu „mate pair” umożliwia ekonomiczną, szybką i precyzyjną identyfikację punktów złamań u pacjentów z BCTs. Dodatkowo, przy *de novo* BCTs identyfikacja punktów złamań niesie ze sobą możliwość diagnozy pacjentów oraz identyfikacji nowych genów/*loci* będących przyczyną choroby u objawowych nosicieli *de novo* BCTs.

2. SUMMARY

Title: Break point mapping in symptomatic balanced chromosomal translocations

Balanced chromosomal translocations (BCTs) are relatively common structural chromosomal rearrangements usually associated with normal phenotype. However, in some cases BCTs may be the cause of the disease. Nevertheless up to 26% of BCTs carriers exhibit some kind of pathogenic phenotype. *De novo* balanced translocations are observed with the frequency of one in every 2000 live births. The higher percentage of symptomatic BCTs carriers versus the general populace suggests a causative link between BCTs and pathogenic phenotype making *de novo* BCTs carriers a natural experimental design for gene disruption and/or deregulation with the breakpoint position directly pointing to the candidate gene(s) and/or regulatory elements critical for observed phenotype. There are at least three molecular mechanisms of by which BCTs may cause abnormal phenotype: (i) direct gene disruption through intragenic breaks may result in gene expression impairment, (ii) breakpoints may intersect open reading frames of genes, and when two genomic segments are joined together after BCT gene fusion is created (also known as a chimeric gene), (iii) chromosomal breakpoints may lead to dislocation and/or disruption of non-coding regulatory elements, which alter the expression of intact genes near the breakpoints, such a phenomenon is called a position effect. Traditionally, BCTs were mapped using conventional methods, such as karyotyping, fluorescence in situ hybridization (FISH), Southern blot, inverse or long-range PCR. However, all these approaches are laborious, time consuming and often provide limited resolution of breakpoint positions. The aim of this project is to identify the new *loci* responsible for the disease in patients with symptomatic *de novo* BCTs. The project did not include patients with:

- Robertsonian translocations.
- BCTs that were inherited from one of the healthy parent.
- aCGH detected copy number variants that were not present in any of the healthy parent.
- The break point of the translocation was present in a centromere or other repetitive sequence that could not be mapped specifically to the reference genome.

In the project 11 patients have been analyzed. To identify the exact structure of the break point shallow mate pair whole genome sequencing with Illumina next generation sequencing technology has been performed. Results were verified with Sanger sequencing. If the translocation did not directly disrupt any gene, qPCR of genes in the damaged topological

domains was carried out. In all 11 analyzed patient the exact structure of the BCTs have been identified. In 5 cases the translocation disrupted genes with undefined or poorly defined impact on the human phenotype. In one case both break points were present in non-coding DNA. In the missing 5 patients genes with known impact on human phenotype were broken. Disrupted genes with unknown impact on the phenotype include: *BAHD1*, *EFNA5*, *PPP2R5E*, *SLC4A10* and *TXNDC5*. Probably all this genes presented above except *TXNDC5* were responsible for the disease of the respective patients. In *TXNDC5* case probably the phenotype is linked to a unverified decreased expression of a gene *TXNDC5* in the damaged domain. The BCTs that was found on both chromosomes in non-coding DNA is responsible for disturbing epigenetic balance of genes in the topological domains leading to an increase in expression of *DKK2* gene, possibly responsible for the mild phenotype of the patient. *EBF3*, *NAA15*, *NIPBL*, *RET*, *SLC6A1* and *TRPS1* are genes with a well-established impact on humans and their disruption is responsible for the disease present in the analyzed patients. One case, in which *ZNF42* has been damaged remains unclear. While mutations in *ZNF423* are inherited in an autosomal dominant way, the phenotype of the patient differs from usual *ZNF423* affected individuals. In conclusion, shallow mate pair whole genome sequencing is a fast, affordable method for detecting the exact structure of BCTs. Additionally we prove that identification of break points in *de novo* BCTs is not only useful for diagnosis of patients but also for detection of new *loci* not yet identified as responsible for disease in humans.

Victor Murua

KIEROWNIK
Zakład Genetyki Medycznej

prof. dr hab. n. med. Rafał Płoski

